

Technical Disclosure Commons

Defensive Publications Series

November 06, 2017

Room noise reduction in audio and video calls

Jan Schejbal

Follow this and additional works at: http://www.tdcommons.org/dpubs_series

Recommended Citation

Schejbal, Jan, "Room noise reduction in audio and video calls", Technical Disclosure Commons, (November 06, 2017)
http://www.tdcommons.org/dpubs_series/785



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

Room noise reduction in audio and video calls

ABSTRACT

Audio or video conferences are often accompanied by room noises, e.g., papers being shuffled, keyboards being tapped, table-tops being tapped, noise from fans, air-conditioning, etc. Although commonly found in various in-room environments, such noises can be annoying and reduce the quality of audio in calls. Techniques of this disclosure reduce noise in microphone pick-up during calls or other in-room recording situations. The techniques utilize optimal placement of microphones, and signal processing techniques such as array processing, speech detection, noise cancellation, etc.

KEYWORDS

- Noise cancellation
- Acoustic array
- Speech detector
- Microphone placement
- Conference room

BACKGROUND

In-call noise can make a call, e.g., a conference call, inharmonious, hard-to-follow, straining, and unpleasant at the far end. Sources of such noise include, e.g., items being shuffled on a tabletop, doors being opened or shut, footsteps as people walk in and out, noise from fans, air-conditioners and other equipment, etc. Although such noises are practically an inevitable component of any meeting, the noises can sometimes constitute a large-enough portion of microphone pick-up that a far-end listener of the call has genuine difficulty in following near-end conversation.

DESCRIPTION

The present disclosure describes techniques to mitigate noise in conference rooms and other in-room recordings.

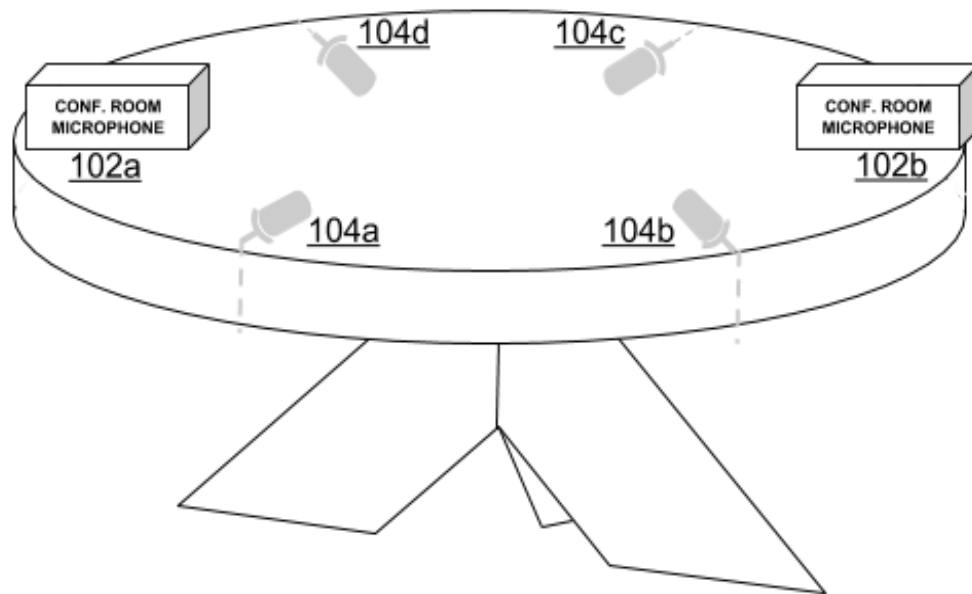


Fig. 1: Conference table with multiple microphones

Fig. 1 illustrates an example of noise cancellation, removal, and/or suppression, per techniques of this disclosure. One or more microphones (102a-b) are placed on a conference table to capture speech, e.g., from participants in a call. As explained previously, the microphones also capture undesirable sound, such as noises of items being shuffled on the table, taps or knocks on the table, etc.

One or more additional microphones (104a-d) are placed on the table such that the additional microphones capture noise. For example, microphones 104a-d can be placed on or firmly connected to the top surface of the table, the bottom surface of the table, embedded within the material of the table, etc. Noise captured by microphones 104a-d is used to cancel

noise from speech-bearing microphones 102a-b, as illustrated in Fig. 2.

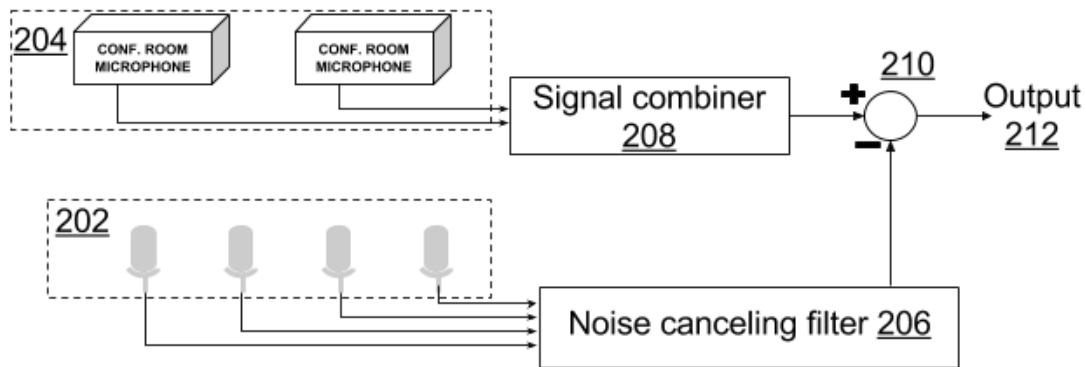


Fig. 2: Noise cancellation

Fig. 2 illustrates noise cancellation via the use of noise-capturing microphones (202). The noise-capturing microphones are close to the sources of noise (e.g., paper-shuffling, tapping on the table, etc.), and hence bear a larger-magnitude version of the noise found in the audio signal captured by speech-bearing microphones (204). Signals from the noise canceling microphones are combined using a noise-canceling filter (206). For example, the filter may linearly combine noise samples across noise-capturing microphones and across windows of time. In addition, noise-canceling filter 206 can have adaptive coefficients to adapt to the changing nature of noise found in the conference room.

In a similar manner, signals from speech-capturing microphones are combined using a signal combiner (208). The output of noise canceling filter closely resembles the noise present at the output of the signal combiner. A subtraction unit (210) subtracts noise filter output from signal combiner output and produces an output (212) that is relatively free of noise.

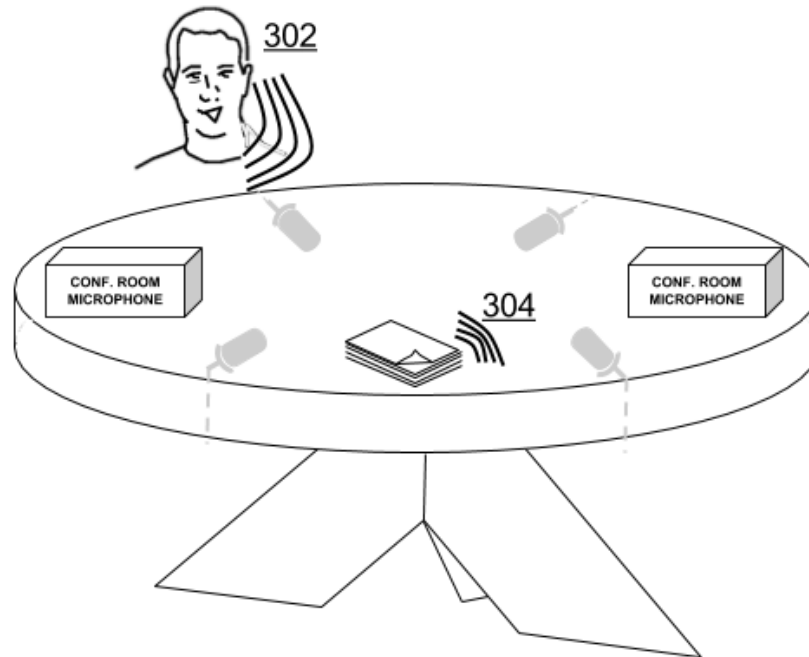


Fig. 3: Separating noise from signal using differential speeds of sound in media

Fig. 3 illustrates the separation of speech from noise using the principle that the speed of sound is different in air than in solid materials, e.g., the material of the conference table. In Fig. 3, the microphones in the room, e.g., those on the table, are calibrated. For example, the microphones are able to determine the source of a sound using knowledge of their relative positions and the time difference between arriving sound waves.

For example, as shown in Fig. 3, as a human speaker (302) speaks, the speech is accompanied by an undesirable paper-shuffling sound (304). The human speech and the paper-shuffling sound are picked up by the microphones. Based on time differences between arrival of the paper-shuffling sound at each of the microphones, a determination is made that the paper-shuffling sound traveled through, and hence likely originated from, the material, e.g., wood, of the table.

Based on the time differences between arrival of the speech of the human speaker at each of the microphones, a determination is made that the speech traveled through air, and hence probably originated from a human speaker. Having identified waveforms corresponding to desirable (e.g., human speech) and undesirable (e.g., paper-shuffling sound) sounds, a signal combiner advantageously coheres speech signals and suppresses noise. For example, a signal combiner can align in time speech signals from different microphones and combine the aligned signals, thereby enhancing speech-signal strength. Similarly, a signal combiner can misalign noise signals in time, thereby diluting such signals.

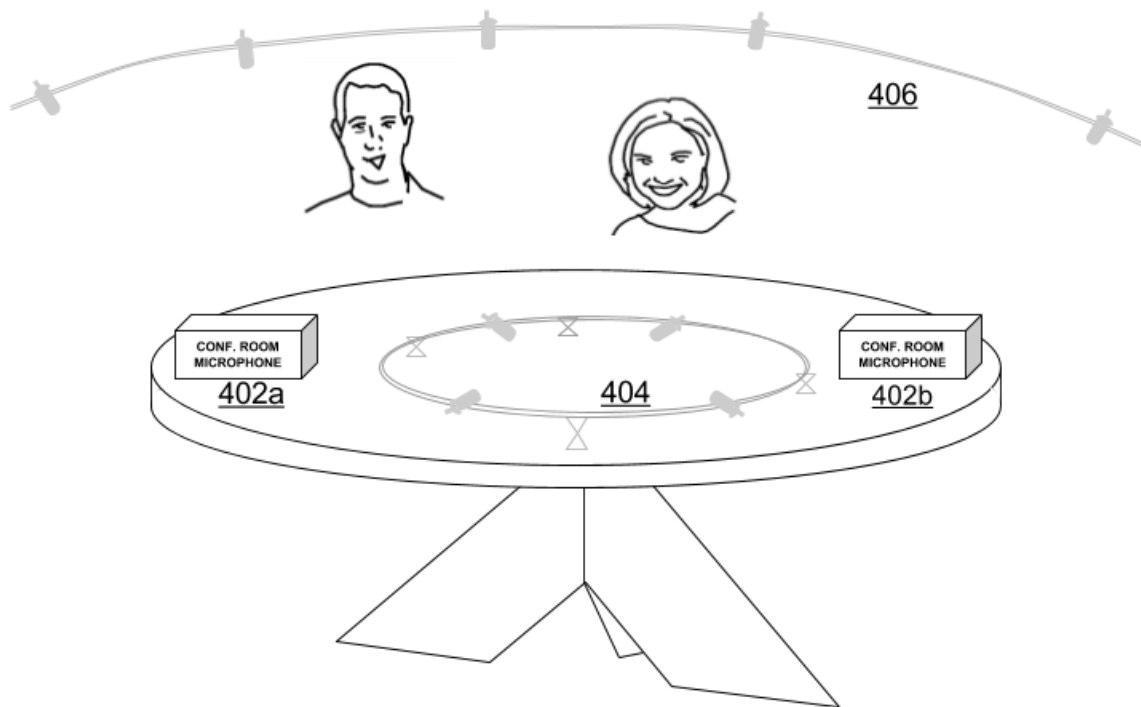


Fig. 4: An array of microphones to suppress below-head level sounds

Fig. 4 illustrates the use of a microphone array to suppress noise. Speech usually originates above table-level, e.g., at head-level of a human speaker, and noise, e.g., paper-shuffling, keyboard-tapping, etc., originates at table level. Techniques of this disclosure make

advantageous use of this observation and use microphone beamforming techniques to reject sounds that originate from below head level.

The microphone array can be installed at a variety of locations, e.g., just above the table surface (404), along a wall of the conference room (406), etc., and can include the conference room microphones (402a-b). Beamforming techniques are used to point the directional gain of the microphone array towards a certain direction, e.g., towards the head-level of the human speakers, and away from certain directions, e.g., away from the space of the conference room that is below head level or at-or-below table level. Although Fig. 4 shows the elements, e.g., microphones, of the microphone array as being connected by wires, wireless microphones can also be used.

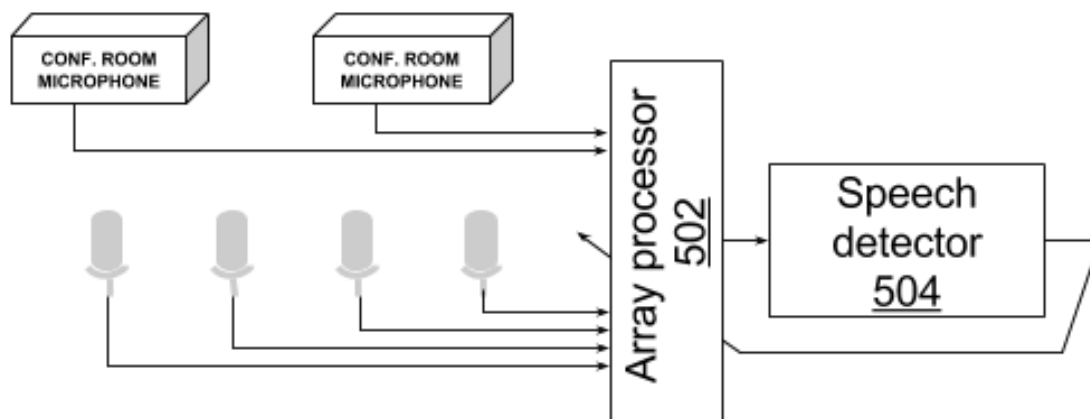


Fig. 5: Using a microphone array to point towards sources of speech

Fig. 5 illustrates the use of a microphone array directed towards sources of speech using an array processor (502) and a speech detector (504). Array processor 502 receives as input a number of microphone signals. The array processor provides directional gain for the microphone array towards a preferred source, e.g., a speech source, by combining microphone signals.

Since speech sources typically change during the course of a meeting, the weights used for combining the microphone signals are adaptive. This is indicated in Fig. 5 by a diagonal arrow that crosses the array processor. The output of the array processor is fed to the speech detector, which detects the presence or absence of speech. If speech is absent, e.g., the array is directionally pointed at a noise source, the speech detector sends a signal to the array processor to re-adapt its weights to point to a speech source. In this manner, a speech detector is used to direct an adaptive array to enhance speech signals and reject noise. Fig. 5 illustrates multiple microphones acting collaboratively, e.g., through an array processor, to home into signals that are detected as speech by a speech detector. However, it is also possible for each microphone to singly act in consort with speech detector in order to home into sounds detected as speech.

Any of the above-described techniques, e.g., microphone placement, array processing, speech-detection, noise cancellation, etc., can be used individually or in combination with other techniques to reduce noise in speech-bearing signals, or reduce the microphone gain (sensitivity) towards noise sources.

CONCLUSION

Techniques of this disclosure reduce noise in microphone pick-up during calls or other in-room recording situations. The techniques utilize optimal placement of microphones, and signal processing techniques such as, e.g., array processing, speech detection, noise cancellation, etc.